# THE INFLUENCE OF SYMBOLIC AND GRAPHICAL LANGUAGE MANIPULATIONS ON ANSWERS TO SELF-ADMINISTERED QUESTIONNAIRES: RESULTS FROM 14 EXPERIMENTAL COMPARISONS

by

Leah M. Christian and Don A. Dillman*

## ABSTRACT

*This paper reports results from 14 experimental comparisons designed to test 7 hypotheses about the effects of two types of nonverbal languages (symbols and graphics) on responses to self-administered questionnaires. The experiments were included in a large-scale survey of 1,042 university students. Significant differences were observed for most comparisons, providing support for all seven hypotheses. These results confirm that respondents' answers to questions in self-administered surveys are influenced by more than words. Thus, the visual presentation of questions must be taken into consideration when designing such surveys and, especially, when comparing results across surveys in which the visual presentation of questions is varied.*

It has been recognized for many years that answers to self-administered questionnaires are influenced by the way in which questions and answers are displayed on questionnaire pages (e.g., Wright and Barnard, 1975, 1978; Rothwell, 1985; Smith, 1993, 1995). However, our scientific understanding of the natures of these effects is not well developed. Although, it has been argued on theoretical grounds that visual layout and design make a difference in how people answer questionnaires (Sless, 1994; Jenkins and Dillman, 1997), more experimental evidence is needed to understand how changing the visual presentation of individual survey questions influences people's answers.

---

* Don A. Dillman is the Thomas S. Foley Distinguished Professor of Government and Public Policy in the Departments of Sociology and Rural Sociology and Deputy Director of the Social and Economic Sciences Research Center (SESRC), and Leah M. Christian is a Research Assistant and graduate student in the Department of Sociology and SESRC at Washington State University, Pullman, Washington 99164-4014. The authors wish to acknowledge the financial support provided by The Agricultural Research Center under Western Region Project W-183, The Social and Economic Sciences Research Center, The National Science Foundation, the USDA-National Agricultural Statistics Service, and The Gallup Organization. Appreciation is also expressed to Brian McQueen for his help with part of this analysis and Thom Allen who served as study director for collection of data analyzed here. Questions should be addressed to dillman@wsu.edu.

In contrast to interviews, which rely mostly on verbal language (or words) for presenting questions to respondents, questions in self-administered questionnaires are presented in nonverbal languages as well. These nonverbal languages include symbolic language (the use of symbols with shared cultural meaning), numeric language (the use of numbers), and graphical language (the use of multiple design features such as font size, brightness, location, and spacing) that may convey certain meanings, apart from those conveyed solely by words.

Our purpose in this paper is to report results focusing on two of these nonverbal languages, symbolic and graphical, which have been given little attention. We have included six independent manipulations and one combined manipulation of symbolic and graphical languages to determine whether answers to survey questions were affected by these manipulations in a paper self-administered questionnaire. All experiments were included in a survey of university students designed to obtain an evaluation of their student experience at Washington State University. This paper seeks to understand how manipulations of symbolic and graphical languages affect respondent answers and to provide general information about the extent to which survey designers must take into account the visual presentation of information in questionnaires.

## THEORETICAL BACKGROUND

It has been argued that the survey process is governed by the conduct of conversation that is used in everyday life and in many instances researchers violate the four assumptions that underlie the conduct of conversation when conducting surveys (Schwarz 1996). These four assumptions or maxims are the maxim of manner (the contribution should be able to be understood by the

audience), maxim of relation (the contribution should be relevant), maxim of quantity (the contribution should be as informative as necessary), and the maxim of quality (the contribution should be true). These assumptions combine to form a cooperative principle whereby one's contribution should be clear, informative, relevant, and truthful (Schwarz 1996, p.11). In the survey process, respondents assume the researcher is following this conduct of conversation so they design their answers in compliance with these maxims. Therefore, the researcher must be careful when designing questionnaires because the respondents assume that all information given to them is relevant to the task at hand, namely completing the questionnaire, and try to make sense of all the information given to them.

Since respondents act as cooperative communicators they will make their best effort to make sense of the survey questions by drawing on all information given by the researcher (Schwarz 1996, p.41). "In a research situation, the researcher's contributions include apparently formal features of the questionnaire, such as the numeric values of rating scales or the scales' graphical layout" (Schwarz 1998, p.182). Especially when respondents are unsure what is being asked or how to report their answer, they will draw information from the context of the conversation, in self-administered questionnaires this includes the context of the question. Schwarz (1996, p. 48) says, "to disambiguate its meaning respondents turn to the context of the question, much as they would in daily life" and this context includes the formal features or nonverbal languages of the questionnaire. Thus, the words make up an important source of question meaning for respondents but so do these formal features or nonverbal languages used in self-administered questionnaires.

Self-administered questionnaires consist of information presented in four distinct languages: verbal, numerical, graphical, and symbolic. These languages can independently and jointly influence respondent behavior. Verbal language, the words of a questionnaire, is used in all modes of data collection and as such is very important in survey design. A great amount of research in survey methodology has focused on the importance of carefully choosing words to convey the researcher's meaning (Schuman and Presser, 1981; Payne, 1951; Sudman and Bradburn, 1974). In self-administered questionnaires, the additional nonverbal languages (numerical, graphical and symbolic) may affect whether questions are read, the order in which they are read, and the meaning conveyed to respondents. Thus, the numerical, graphical, and symbolic languages influence how respondents interpret the verbal language of a self-administered questionnaire.

Research has shown that numerical language independently influences respondent behavior (Schwarz, et. al. 1985, 1991). Schwarz, et. al. (1985) have shown that respondents gain information about the researcher's expectations from the response alternatives used in surveys. Specifically, multiple experiments by Schwarz and others have shown how respondents use the verbal and numeric labels of ordinal scales as a frame of reference in determining their own answer. Schwarz, et. al. (1991) have shown that manipulating the numeric values attached to the endpoints of opinion scales affected the distribution of responses. They conducted multiple experiments with a university and adult German population comparing linear scales with endpoints numerically labeled 0 to 10 with scale endpoints labeled –5 to 5 (keeping the verbal labels attached to the endpoints constant), and found that this independent numerical language change affected respondents answers (Schwarz, et. al. 1991). They argue that respondents

"disambiguate" the meaning of the verbal labels by using the numeric values attached to those labels. Thus, the respondent uses the numerical language as a source of question meaning above and beyond the verbal labels of the scale.

Graphical language is used to summarize features such as of size, brightness and color, shape, location and spatial arrangement. Graphical language acts like paralanguage, optional voice effects that accompany the sounds of an utterance and may convey meaning, used in aural communication. Thus, graphical language includes features of the questionnaire that accompany words, numbers, and symbols and convey meaning. The verbal, numerical and symbolic languages "can only be transmitted through the visual channel via graphic paralanguage" (Redline and Dillman 2002, p.181). Thus, graphical language is the visual conduit through which all other languages of a questionnaire are expressed. For example, when a respondent reads a question, the graphical language can vary the stimulus seen by the respondent by changing the size of words, the color of symbols, or the graphical layout of a scale and these changes can influence respondent answers.

Tom Smith (1993, 1995) reviewed tests of three independent accidental graphical manipulations where the graphical location of answer spaces, the size of answer spaces on an open-ended question, and the size of boxes used to display a socioeconomic scale were all varied. In the first manipulation, a font error misaligned the response categories in the second item of a three item question. In this version, the yes box was placed in the no column and the no box was further out to the right. Analyzing the response distributions, it is apparent that some respondents confused the yes and no boxes because the graphical location of the answer was affecting how

respondents answered. Secondly, Smith recognized that when respondents were given a larger answer space for an open-ended question they wrote longer answers. Lastly, he reports that changing the size of the answer boxes from a stack of ten vertical boxes of equal size to a stack of ten vertical boxes of different sizes reflecting a pyramid shape where the bottom boxes were larger than the top boxes resulted in more respondents selecting the lower, larger width boxes on the pyramid version because they thought the wider width indicated that more people were actually in those categories. Thus, the respondent gained information about how the researcher thought socioeconomic status was distributed from the graphical layout of the scale. Each of these examples shows that independent manipulations of one feature of graphical language can influence how people respond to survey questions.

Schwarz (1998) replicated the Smith experiment above on socioeconomic status by including a question about academic performance on a survey given to a university population where some respondents were given the ten vertically stacked boxes, some were given the pyramid version, and some were given a new test where the boxes replicated an onion shape where the middle boxes were larger and the top and bottom boxes were smaller. These results replicated Smith's findings that respondents rated their academic performance less favorably on the pyramid version (more respondents chose the wider boxes). In addition, the mean for the onion version was similar to that of the stacked boxes indicating perhaps that respondents assume an onion shaped distribution even when that is not communicated through the size of the boxes (ie. in the stacked box version). Thus, respondents were gaining information about researcher's expectations by the graphical layout or representation of the scale.

Schwarz and Hippler (Schwarz et. al. 1996) also conducted a study manipulating one aspect of graphical language by grouping two questions (asking about general satisfaction and marital satisfaction) together in one single black box, and they found that respondents were more likely to see the two questions as related. However, when the questions were separated, each in their own black box, respondents treated them as separate questions. The results show that when respondents see the questions as related the correlation between the two questions is reduced because respondents evaluate marital satisfaction and general satisfaction as two distinct questions and do not include information about marital satisfaction in the general satisfaction question because they feel they have already reported that information and are assuming the researcher is following the maxim of relation so they would not be asked to give redundant information.

Symbolic language uses signs and symbols that have cultural meaning to convey that meaning to respondents. For example, an arrow is used to tell a person to go in the direction the arrow is pointing. Current experimental evidence shows that combined symbolic, graphical, and verbal language changes reduced respondent branching errors significantly. In a student classroom experiment in which 1,266 students completed one of three versions of a questionnaire, commission errors (not skipping ahead when directed to do so) were reduced from 20% to between 7-9% for the newly designed instructions (Redline and Dillman, 2002). In a follow-up test imbedded in the 2000 Decennial Census of the United States, commission errors were reduced from 21% to 13-15% for two similar methods (Redline, Dillman, Dajani, and Scaggs, Forthcoming). A shortcoming of both of these experiments is that the independent effects of the verbal, symbolic, and graphical language changes were impossible to disentangle. These

experiments have shown that the use of visual design can make a significant difference in whether respondents follow branching instructions. However, they did not reveal which specific languages were responsible for the changes in respondent behaviors and only two discrete examples of combined manipulations were provided.

By conducting a series of experiments involving individual manipulations of graphical and symbolic language, we attempt here to provide insight into a range of potential response effects from these individual changes. In addition, one combined symbolic and graphical language manipulation will illustrate how the languages can be used to work in concert with each other. Since respondents use the languages of the questionnaire to extract information about the researcher's expectations, it is important for the four languages to be giving consistent and not conflicting information. In contrast to the work mentioned in the preceding paragraph, most of these tests involve individual questions that all respondents were expected to answer. Thus, this paper is an effort to further move beyond branching instructions as a focus of language effects, as well as further investigate how manipulations of graphical and symbolic languages affect response behavior using self-administered questionnaires.

## GRAPHICAL AND SYMBOLIC LANGUAGE MANIPULATIONS

### Graphical language manipulations

This paper summarizes four independent graphical language manipulations to determine the effects on response behavior. In the first manipulation, graphical location was varied by placing a special instruction before and after the response categories (Figure 1). The location before the response categories places the instruction in the navigational path so respondents can process the

instruction after reading the query but before a response is required. It has been argued that an important goal of questionnaire construction is to use graphical language in a way that makes the elements (number, query, instructions, and answer choices) of a question appear as a distinct group and instructions should be provided exactly where they are needed by the respondents (Dillman, 2000, p. 96-99). In this experiment a Yes/No question was developed that was expected not to apply to a number of respondents.[1] In one version, the special instruction to skip to the next question when applicable was placed ahead of the response categories, and in the other version placed immediately after the categories. It was expected that more people would skip the question when the instructions were located before the response options because they were more likely to see and thus read the instructions before answering. Therefore, the graphical location of the instructions would influence whether the respondents read them, subsequently influencing whether they chose to skip or answer the question. Thus, it is hypothesized that special instructions are more likely to be followed correctly if they are located in the navigational path just before, rather than just after, the place of their intended use.

Another graphical manipulation changed two linear ordinal scales (all scale points listed vertically) to a nonlinear layout (where responses were triple and double banked). Responding to ordinal scales requires a different mental process than responding to nominal scales. Using an ordinal scale, the respondent must decide where to place oneself on an implied continuum, whereas responding to nominal scales requires respondents to compare categories in order to select the best one. Graphically speaking, a linear layout of ordinal scales would seem to facilitate the process of identifying where a respondent best fits on the continuum. When the

[1] A filter question was not used because this experiment was specifically testing the location of special instructions within or outside of the navigational path.

responses are viewed in a nonlinear layout, the graphical language conveying the four or five point scale is interrupted. Nonlinear layouts such as double and triple-banking responses are commonly used to save space without the implications on responses considered. This use might also be problematic given that some respondents read horizontally and others vertically and potential confusion could result especially when the layout orients respondents to process the questionnaire vertically (as in this case). For this manipulation, two experiments were conducted, linear layout versus triple banking of an ordinal scale and linear layout versus double banking of an ordinal scale (Figure 1). It is hypothesized that the percentage of respondents choosing categories from the top line will be significantly greater than from the bottom line, and in particular the category just to the right of the first category will increase in the nonlinear layout because some respondents may read the top line only, ignoring the bottom line.

Graphical location and spatial arrangement can also be manipulated by increasing the space between response categories. No rules appear to exist for how far apart response categories should be placed from one another. Intuitively, it would seem that response categories should be placed equal distances from one another. However, the ramifications of unequal distances between categories are still unclear. As mentioned earlier Schwarz (1996) reports that respondents use the response categories to gain information about how to report their answers. In one experiment, we are testing a nominal scale where respondents are expected to gain a lot of information about what the question is asking from the response categories (Figure 1). In the second experiment, we are testing an ordinal scale where respondents simply place themselves in the correct category but know that the question is asking for a percentage from the query. In these manipulations, we are testing whether increased space between some response categories

and the other categories will increase the visual prominence of that category and result in more respondents choosing that category.

*(I shortened the following paragraph and combined it with what we expect – what do you think?)* It was found in a branching instruction experiment (Dillman and Carley-Baxter, 1999) that 3 out of 52 questions produced significantly different responses among treatment groups. Two of the three questions (one using a nominal scale and the other an ordinal scale) produced significant results because of the greater selection of one response option when it was graphically separated from the other options (by increasing the space between it and the other options). In the versions where these responses were chosen more often, answer boxes were placed to the right (instead of the left) of response options. The possibility that this change in answer box location influenced the difference in some unknown way suggested the need for testing the spacing of the response options independently of the placement of the answer boxes. These two question formats have been retested in this study one testing a nominal scale about life goals and the other testing an ordinal scale about recreational center usage (Figure 1). It is hypothesized that more people will choose the answers to the substantive categories that are given greater visual prominence by being spatially set off from the others, especially in the nominal scale question because respondents need the response categories to understand what the researcher is asking.

The final independent graphical manipulation, involved three experiments testing how increasing the size of open-ended answer spaces would affect respondent behavior. The size of the open-ended answer space was doubled on one version of each of the three experiments (Figure 1). As discussed earlier, size conveys meaning about the researcher's expectations and has been shown to affect response behavior. In open-ended questions size may become more important as

respondents do not have response categories or scales to help determine what the researcher would like for them to report. In self-administered questionnaires where the query and response space are the only information given to the respondents and further probing by interviewers cannot be done, it has been shown that respondents typically provide shorter, less complete answers to open-ended questions than when surveyed in interviews (Dillman 2000). Based upon work by Stember (1956), Smith (1993) has argued that allowing more space for recording open-ended answers in interview-administered surveys produces longer recorded responses that may be closer to actual verbatim. It is hypothesized that providing a larger answer space will produce longer answers that contain more substantive themes.

**Symbolic language manipulations**

The first symbolic language manipulation tested the effects of adding lines in open-ended answer spaces. Since it is hypothesized that the amount of space conveys meaning to respondents, two questions were tested each with a different size answer space. In addition, the substantive topic of the question was varied for each question. For example, the first question tested here asks what additional stores or businesses the respondent would like to see in the area (see Figure 1). Here, the answer space was small and adding lines only gave the respondent two lines on which to write, suggesting that only a few words and stores should be used when answering. The second question asked the respondent to report one or two changes that would improve the educational experience at the university. Here the respondent was provided a much larger space and the lines version provided twelve lines. In this case, the presence of the lines in a large space might make the space appear larger increasing item nonresponse on that version. In addition, it is expected that when given a larger answer space, respondents will report the answers in

paragraph format and will use the lines when provided. Thus, it is hypothesized that the addition of lines will produce less words and themes on the first question and will increase the perceived cognitive burden on the second question with a larger space, leading to greater item nonresponse.

The second symbolic manipulation was the addition of an arrow to direct respondents towards answering a subordinate question. Sometimes questionnaire designers wish to direct respondents who choose a particular answer to a subordinate question for which another answer is desired. An example, and the one tested here, was imbedded in a question that asked where the respondent wanted to live after completing college, "Eastern Washington" or "Somewhere Else" (as shown in Figure 1). After the answer choice "Somewhere Else", the word "Where" was listed on the same line as the answer choice, approximately 10 spaces beyond the end of the category description and 26 spaces beyond the answer box. In one version, an arrow was placed between the answer "Somewhere Else" and the subordinate question "Where."

It has been shown that when respondents read text they focus on a space of about 2 degrees, or 8-10 characters in width (Kahneman, 1973). This distance is known as the *foveal view*. Thus, the arrow was placed so the subordinate question would be visible in the foveal view. An arrow is a symbol that is culturally defined to focus one's attention in the direction the arrow is pointing. It was reasoned that respondents were more likely to see and, as a result, respond to the subordinate question when an arrow was placed between the category description and the subordinate question. This is a test of the independent effect of symbolic language on item nonresponse and the arrow was chosen because it was used in a combined manipulation of multiple languages that reduced branching errors in a study by Redline, et. al. (Forthcoming). It is hypothesized that the

addition of an arrow between a response category and a subordinate question will increase the likelihood that respondents will answer the subordinate question.

**Combined graphical and symbolic language manipulation**

The final manipulation was a combined graphical and symbolic manipulation where a polar point scale (labeled endpoints only) was compared to an answer box (where respondents write-in the number corresponding to their answer). This experiment tests the effects of eliminating graphical (linear layout of choices) and symbolic language (the check boxes associated with each category) on one version (see Figure 1). It has been argued that the use of an answer box might make it possible to provide equivalent stimuli across survey modes because the additional graphical and symbolic language cannot be provided in aural modes (Dillman, 2000, pages 235-236). Determining the effects of removing these languages within self-administered questionnaires, and thus relying solely on words, is a related concern that needs to be tested before further tests across modes are conducted. Three items tested the effect of answer boxes vs. linear polar point scales for responding to ordinal scales.

One of the potential difficulties of the number box format is that it requires respondents to remember the specifics of the scale when providing their answer, whereas the inclusion of graphical and numerical information in the polar point scales provides a reminder of how the scale is constructed and is located within the foveal view (8-10 characters) of the response options. As discussed earlier, experiments by Smith (1993, 1995) and Schwarz and Hippler (Schwarz 1996) show that changing the graphical layout of a scale can influence respondent behavior. Since the graphical layout of the scale provides additional support to the verbal and

numeric languages in the polar point version, it is hypothesized that the answer box will produce different responses since the additional graphical and symbolic language support is removed.

## PROCEDURES

The 18 experiments were embedded on pages 2, 3, and 4 of a four-page questionnaire developed for assessing the student experience at Washington State University and conducted from March to April 2001. It was printed in a two-column format on 8-1/2 x 11 inch pages, with a colored background being used to contrast with white answer spaces provided for both open- and closed-ended questions.

Four versions of the questionnaire were mailed to equal subsamples (450) of a random sample of 1,800 undergraduate students living in the Pullman, Washington area (students on other campuses or enrolled in the distance degree program were excluded). A $2 incentive was enclosed with the first mailing. A follow-up postcard and one replacement questionnaire were mailed, obtaining a response rate of 57.9% (1,042) of the 1,800 questionnaires mailed.

The experimental questions reported here were the same in two of the four versions of the questionnaires[2]. Inclusion of so many experiments in one questionnaire raises issues of whether some of the experiments may have affected results for others. That possibility cannot be ruled out. For the most part, the alternative treatments being tested here are not ones that seem likely to influence respondents in such a way that answers to subsequent questions would somehow be

[2] Four versions were needed for three other experiments involving the construction of pages one and three and have been reported elsewhere (Sawyer and Dillman, 2002). All of the experiments reported here were the same on Versions A and D and Versions B and C.

affected. It is quite possible that percentage distributions to individual questions were influenced through order effects, but we were unable to find evidence that such effects might have impacted the kinds of hypotheses being tested here.

Statistical tests made to evaluate the hypotheses include chi-square tests for differences and t-tests for mean differences, where appropriate. The tests conducted vary depending upon the questions and therefore will be reported for each hypothesis.

## FINDINGS

### Graphical language manipulations

*Graphical location.* This graphical location change dramatically affected whether respondents used the special instructions. When the special instruction "If you haven't had many one-on-one meetings, just skip to Question 9" was placed after the response options, 54.9% of the respondents said "Yes", 40.3% said "No", and 4.8% provided no answer (Table 1). However, when these instructions were placed before the responses, 54.7% said "Yes", 19.1% said "No", and 26.2% provided no answer ($\chi^2 = 116.99$, p = .000 - Yes vs. no vs. Missing by Version). Thus, when special instructions were graphically placed where respondents were more likely to see them, prior to answering the question, the location influenced their decision to answer or not answer the question. Not only did placing the instructions after the responses, produce different results on this question but it also introduced confusion as some respondents used the instructions for the following question. Item nonresponse to the following question increased from 2.6% to 11.0% when the instructions were placed after the responses. It appears that

respondents may have mentally grouped the special instruction with the following question and interpreted the instructions as directing them to skip over it to Question 10.

*Linear vs. Nonlinear scale layout*. Changing the scale layout from a linear to nonlinear layout, affected respondent behavior in one of the two experiments. Respondents were significantly more likely ($\chi^2 = 6.66$, $p = .010$) to select responses from the top line in the nonlinear version (see Table 1). Specifically, 40.4% of respondents chose "Good" and 42.4% chose "Very Good" on the nonlinear version. Whereas in the linear format, 31.3% of respondents chose "Good" and 48.8% chose "Very Good". This difference by version between "Good" and "Very Good" is significant ($\chi^2 = 8.2$, $p = .004$). More respondents chose "Good" when the scale was triple-banked and less respondents chose "Very Good", suggesting perhaps that some respondents focused only on the top line ignoring the bottom line. A slight trend in the same direction exists for the second question, which used double-banking, but the chi-square (top vs. bottom line $\chi^2 = .16$, $p = .685$) nor individual chi-squares for differences between individual answer categories were significant. On this scale, the two horizontal categories "Very Satisfied" and "Somewhat Dissatisfied" do not display the sense of a complete scale, as might be inferred from the three horizontal categories – "Excellent, Good, Poor" – on the question for which differences are statistically significant so respondents may be more likely to read all the options to make sense of the scale. Thus, differences based on content (verbal language) and number of response options between the two tests of this hypothesis could have produced the difference in results.

*Equal vs. unequal spacing between response categories.* The results from testing the effect of equal versus unequal spacing between response categories were significant ($\chi^2 = 4.8$, $p = .028$ -

unequally spaced response vs. all others) on the nominal scale question but not the ordinal scale question ($\chi^2 = 1.4$, p = .844 overall chi-square) (Table 1). In the question where significant effects were found, the content of the response options were more likely to help respondents gain information about what the researcher was asking and as such susceptible to context effects. This resulted in the response option "To have a life partner with whom you have a satisfying relationship" to be selected significantly more times on the unequally spaced version (37.6% vs. 30.8%) when the answer box for this response was separated by two spaces from the other responses. However, in the ordinal scale question, the respondent already knows what the researcher is looking for (a percentage as stated in the query) so the respondent must simply choose the category that fits the percentage they are thinking. We suspect this question was less susceptible to graphical spacing effects because less meaning is gained from the response categories than in the nominal scale question.

*Use of larger answer spaces on open-ended questions*. Varying the amount of answer space on open-ended questions results in the amount of space influencing both the number of words and the number of themes provided in respondents' answers (Table 1). The number of words was hand-counted for each open-ended response and themes were counted as the number of topics mentioned. The coding of themes was completed by one researcher and 10% were verified by another researcher with 90% agreement. The results show that a larger answer space produced longer answers with a greater number of themes. On all three questions, the larger space produced longer answers with a significantly greater number of words 13.3 vs. 9.7 (t=6.5, p=.000), 12.9 vs. 6.6 (t=12.9, p=.000), and 12.0 vs. 10.2 (t=1.8, p=.039). On two of the three questions, the larger answer space significantly produced a greater number of themes or topics

mentioned in the answer, 2.0 vs. 1.8 (t=2.7, p=.003), 2.1 vs. 1.7 (t=8.0, p=.000), but in the third answer the results were in the expected direction but not significant (1.5 vs. 1.4 with t=.7, p=.195). These results confirm earlier findings that larger answer spaces for open-ended questions produce longer answers and extends previous findings by looking at the number of themes mentioned and finds that in two of the three questions, a larger answer space significantly produced a greater number of themes.

**Symbolic language manipulations**

*Addition of lines on open-ended answer spaces.* Virtually all respondents understood that they were supposed to use the lines on that version (only 1.2% of respondents did not use the lines on the first question and .8% on the second question). Thus, respondents knew the cultural meaning attached to lines was to use them when providing their answer. On the first question, a significantly greater number of words (6.7 vs. 5.3, t=-4.91, p=.000) were used on the no lines version suggesting that the addition of the lines shortened the appearance of the answer space encouraging respondents to provide shorter answers. However, the difference between means (2.3 vs. 2.4) was not significant (t=-.62, p=.269) suggesting that instead of reducing the number of stores listed, respondents abbreviated or shortened the name of the stores or businesses resulting in less words used. On the second question, neither the number of words or number of themes was significant across the two versions. However, the addition of lines did increase the number who left the space blank on this question from 19.4% to 26.6% ($\chi^2 = 7.8$, p = .005) (Table 2). However, further analysis indicates that when examining the responses, some respondents did not provide an answer to the question but wrote responses like “None” “I have had a wonderful experience at WSU” suggesting that they felt a need to fill the space even when

not answering the question. This was significantly more likely on the version without lines (7.1 vs 3.1 t=7.06, p=.008). These results show that the effects of adding lines on open-ended answer spaces still needs further research on questions with various sizes of open-ended spaces but differences in the length of answers as well as affects on item nonresponse may occur..

*Addition of arrow to identify subordinate question*. The addition of the arrow to direct respondents towards a subordinate question significantly increased the percentage of eligible respondents answering the subordinate question (93.9 vs 90.8%) (Table 2). However, one concern about the use of the arrow is that it also increased the mentions from noneligible from .5 to 2.1%. In addition, the arrow decreased the number of missing or blank responses (4.2 vs. 8.8%). The overall chi-square is significant ($\chi^2 = 11.65$, $p = .003$). Thus, this independent symbolic language manipulation significantly influenced respondent behavior. The results of this experiment indicate that the use of the arrow did help direct respondents towards a subordinate question and increased the item response to that question; however, the researcher must weigh the adverse effects of adding the arrow as it may direct noneligible respondents to the subordinate question as well.

**Combined Graphical and Symbolic Language Manipulation**

*Polar point scale vs. answer box*. This final experiment tested the combined manipulation of graphical and symbolic language by replacing the graphical layout of a polar point scale with an answer box where the respondent is supposed to write-in the number corresponding to their answer. This combined manipulation produced dramatic differences in the responses given by respondents. First, the use of the number box significantly increased the mean for each of the

questions tested (Table 3). The means changed from 2.4 to 2.8 for the first two questions and from 2.7 to 2.9 for the third question (p=.000 for all three mean differences).

To test for why these increases in means occurred, additional coding was conducted to see if respondents became confused and did not remember the direction of the scale (from positive to negative, consistent with other items in the same section). Responses to each of the three answer box questions were coded based on whether there was any evidence of respondents changing answers and as a control evidence of changing answers on the polar point version were coded as well. To be included in the analysis, both the original and final answers must have been "legible" so any nonlegible answers were not coded. On the answer box version 10% (versus 1% on the polar-point scales) of respondents scratched out answers to at least one of the questions and provided a different answer. A total of 74 respondents made 86 changes in their answers. Most of these errors occurred because respondents reversed the scale on the answer box version: 44 respondents changed from 4 to 2, 10 respondents changed from 5 to 1, 4 changed from 2 to 4, and 1 changed from 1 to 5. These data suggest that removing the graphical layout of the scale from the polar point version by changing it to an answer box introduced confusion in respondents' understanding of the scale. These data suggest that other respondents may have made this error without catching or changing them and this is reflected in the larger means on the answer box version.

To additionally test for respondent confusion, individual correlations were calculated between both versions of these three questions and 13 other questionnaire items about satisfaction with classroom experience that for the most part were not varied across experimental treatments (1

was varied). Each of these items was expected to correlate positively with answers to these three test questions. If confusion existed among respondents who did not change their answers, we would expect lower correlations between the three answer box questions and the 13 other items than for the polar point versions. As expected, all 13 correlations for each of the items using both scalar formats were positive; however the mean for the polar point format was .24 compared to only .14 for the answer box format. In only 4 of the 36 instances was the correlation between the 13 items and the test items higher for the answer box version. It was also observed that the mean difference for the first item (.16) was higher than the mean for the second (.10) with the third item reflecting the lowest mean difference (.03). This was consistent with the expectation that respondents are less likely to make errors on the later items because they become accustomed to using the scale after they have to apply it more times on the same type of question.

**DISCUSSION AND CONCLUSIONS**

Most previous work in survey methodology has focused on question wording as the sole conduit of question meaning. In this paper, we have proposed that the use of nonverbal languages, numerical, symbolic, and graphical, communicate additional meaning to the respondent and can independently and jointly influence respondent behavior. Specifically, we have demonstrated how individual and combined manipulations of graphical and symbolic languages influence respondent behavior.

These 14 experimental tests of 7 hypotheses combine to form an initial analysis of how graphical and symbolic languages influence respondent behavior in self-administered questionnaires.

Significant differences were found for all of the hypotheses (though not on every experimental item) and most of the remaining tests were in the expected direction. Our general conclusion is that the visual design of questions on self-administered questionnaires significantly impacts respondent behavior. In addition, these findings have implications for survey designers when designing self-administered questionnaires as well as surveys that are to be used across modes and for users of survey data to explicitly consider visual design because similarly worded questions may be presented in visually dissimilar ways, resulting in answers not being directly comparable.

Some of the specific findings have straightforward implications for survey design. Special instructions should be placed in the navigational path where they are needed and should be specifically grouped with the proper question because incorrect placement produces consequences for answering that question and whether respondents correctly advance to the next question. The result of adding an arrow between an answer choice and a subordinate question placed just outside the foveal view (8 to 10 characters), show that such symbols can independently guide respondents to appropriate follow-up questions. These findings in conjunction with the findings reported earlier (Redline, et. al., Forthcoming) demonstrate that individual and combined manipulations of visual languages influence how respondents navigate through a self-administered questionnaire as well as influence answers to specific survey questions.

Space limitations in survey questionnaires are often a concern when designing surveys; however, results from this research suggest that space conveys meaning to respondents so care should be

taken when designing surveys. Larger answers spaces for open-ended questions produced longer answers with more themes. The addition of lines may make the space appear larger or smaller depending on what the question is asking. The stores question suggests that when respondents are given too small a space they may shorten their answers or abbreviate things to fit their answers in the space. In addition, item nonresponse increased on a question with a larger space, regardless of whether lines were used and the number of respondents who left the question blank was significantly higher on the version with lines. These two experiments suggest that the amount of space for open-ended answers should be chosen carefully such that respondents are given enough space to answer the question (according to the researcher's expectations) but not too much so that it increases the perceived task difficulty thereby increasing item nonresponse.

Spacing also conveys meaning in closed questions with regards to how response options are spaced in ordinal and nominal scales. The common practice of double and triple banking response options should be done with care as one experimental test found that triple-banking the response options resulted in significant changes in the relative proportions of respondents choosing items from the top line. Thus, double and triple-banking responses could encourage respondents to process horizontally or vertically so that respondents are not receiving equal stimulus when processing the response options. A linear layout of responses facilitates the process of respondents finding where one fits on the continuum in an ordinal scale. Lastly, space between response options was found to produce different results on one of the two experimental tests. Here, the nominal scale encouraged greater dependency on response options to understand what the researcher was asking and as such was subject to context effects. Separating one response by two spaces between the answer box and all other answer boxes significantly

increased its likelihood of selection. These findings combine to show that graphical space in questionnaires communicates meaning to respondents so care should be taken when designing the layout of individual survey questions.

Although the previous experiments reported here focused on individual manipulations of symbolic or graphical language, one experiment simultaneously manipulated the graphical and symbolic languages. This experiment, the use of an answer box vs. a linear polar-point scale, suggests the need to consider the supportive role that symbols and graphics play in conveying question meaning to survey respondents. The removal of the graphical layout of the linear polar point scale at the time of response required respondents to rely solely on the question stem (query) to guide them while answering and resulted in increased means as confusion was introduced because respondents had trouble remembering which end of the scale was positive and which was negative. The answer box version resulted in more changed answers (because of reversals of the scale) and lower correlations on the answer box version versus other satisfaction measures than the polar point version. The direct implication is that when opinion scale questions in self-administered questionnaires are designed with answer formats that require applying information from the stem of the question, answers are likely to be less accurate; therefore, providing a reminder of that information in the response scale can encourage more accurate answers. As mentioned earlier, combined manipulations do not allow us to unpack the individual effects of such changes but do suggest the supportive role that visual languages can play in providing an equal stimulus to respondents such that when meaning is derived from the verbal and nonverbal languages, accidental changes in the nonverbal languages will not influence substantive answers to survey questions.

The hypotheses evaluated here are only a few of many that might be tested. Only two symbolic language changes were evaluated (use of arrows and lines) and many other experiments could be conducted testing other symbols. Although four independent graphical changes – ranging from size of the open-ended answer space to unequal distance between answer choices – additional manipulations of graphical languages need to be tested in other experiments to fully understand the effects of all aspects of graphical language on respondent behavior.

Based on the experimental evidence presented here and in previous writings (Redline and Dillman, 2002; Redline et al., Forthcoming) as well as additional studies discussed earlier (Smith 1993, 1995; Schwarz 1986, 1991, 1998; Schwarz et. al. 1996), it is apparent that survey questions on self-administered questionnaires consist of much more than words. Respondents following the conduct of conversation actively construct meaning as they complete survey questions and gain meaning from the context of the questions which includes the formal features of questionnaire design – symbols, numbers, and graphics. Future research needs not only to test the current hypotheses on other populations using different substantive items, but also needs to evaluate other manipulations, e.g., the role of numbers in getting people to answer questions in a prescribed order, the role of consistency in the use and display of symbols throughout questionnaires, effects of figure/ground variations on whether information gets processed by respondents, and how color and contrast can affect question comprehension. This additional research will aid in developing general principles about how questions should appear on questionnaire pages.

In addition, all of the current research on visual design has focused on mail self-administered questionnaires. The increasing use of internet surveys demands that these ideas be tested on web questionnaires as well as results compared across mail and internet to understand how visual design principles work in different self-administered modes. Furthermore, mixed-mode comparisons are necessary to understand these visual principles as they impact comparisons across self-administered and aural modes (telephone and face-to-face interviews). The recent trend towards greater use of mixed-mode surveys in which researchers attempt to survey some members of a population via one mode, and some by another, suggest that a priority for future research is to understand the fundamental causes of such modal differences. Such work is essential for taking us beyond our current stage of understanding survey responses by verbal language alone, to beginning an understanding of how verbal and nonverbal languages work together to convey meaning to respondents.

**REFERENCES**


Dillman, Don A. (2000) *Mail and Internet Surveys: The Tailored Design Method.* 2nd Edition. New York: John Wiley and Sons, Inc.

Dillman, Don A. and Lisa Carley-Baxter. (1999) Unpublished data. Social and Economic Sciences Research Center, Washington State University, Pullman, WA.

Jenkins, Cleo R. and Dillman, Don A. (1997) "Towards a Theory of Self-Administered Questionnaire Design." In Lyberg, et. al. (eds.), *Survey Measurement and Process Quality*. New York: John Wiley and Sons, Inc.

Kahneman, D. (1973) *Attention and Effort*. Englewood Cliffs, NJ: Prentice Hall.

Payne, S. L. (1951). *The art of asking questions*. Princeton, NJ: Princeton University Press.

Redline, Cleo, Don A. Dillman, Aref Dajani, and Mary Ann Scaggs. (Forthcoming) "Navigational Performance in Census 2000: An Experiment of the Alteration of Visually Administered Languages." *Journal of Official Statistics.*

Redline, Cleo and Don A. Dillman. (2002) "The Influence of Alternative Visual Designs on Respondents' Performances with Branching Instructions in Self-Administered Questionnaires." In, R. Groves, D. Dillman, J Eltinge, and R. Little (eds), *Survey Nonresponse*. New York: John Wiley and Sons, Inc.

Rothwell, Naomi D. (1985) "Laboratory and Field Response Research Studies for the 1980 Census of Population in the United States." *Journal of Official Statistics*, 1 (2): 137-157

Sawyer, Scott and Don A. Dillman. (2002) "How Graphical, Numerical, and Verbal Languages Affect the Completion of the Gallup Q-12 on Self-Administered Questionnaires: Results form 22 Cognitive Interviews and a Field Experiment." Social and Economic Sciences Research Center Technical Report 02-26, Pullman, WA.

Schuman, Howard, and Stanley Presser. (1981). *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. New York, NY: Academic Press.

Schwarz, Norbert, Carla E. Grayson, and Bärbel Knäuper. (1998). "Formal Features of Rating Scales and The Interpretation of Question Meaning." *International Journal of Public Opinion Research*, 10(2): 177-183.

Schwarz, Norbert. (1996). *Cognition and Communication: Judgmental Biases, Research Methods, and The Logic of Conversation.* Mahwah, NS: Lawrence Erlbaum.

Schwarz, Norbert, Bärbel Knäuper, Hans J. Hippler, Elisabeth Noelle-Neumann, and Leslie Clarak. (1991). "Rating Scales: Numeric Values May Change the Meaning of Scale Labels." *Public Opinion Quarterly,* 55: 570-82.

Schwarz, N., H.J. Hippler, B. Deutsch, and F. Strack. (1985) "Response scales: Effects of category range on reported behavior and subsequent judgments." *Public Opinion Quarterly*, 49:388-395.

Sless, David. (1994) "Public Forums: Designing and Evaluating Forms in Larger Organizations." *Proceedings of Public Graphics,* pp. 9.2-9.25. Lunteren, The Netherlands.

Smith, Tom W. (1993) "Little Things Matter: A Sampler of How Differences in Questionnaire Format Can Affect Survey Responses." GSS Methodological Report No. 78. National Opinion Research Center, Chicago, IL.

Smith, Tom W. (1995) "Little Things Matter: A Sample of How Differences in Questionnaire Format Can Affect Survey Responses." Paper presented at the American Association for Public Opinion Research, Ft. Lauderdale, FL: May.

Sudman, S. and Bradburn, N. (1974). *Response effects in surveys: A review and synthesis*. Chicago: Aldine.

Stember, Charles Herbert. (1956) "The Effect of Field Procedures on Public Opinion Data." Unpublished Ph.D. dissertation, Columbia University.

Wright, P. and P. Barnard. (1975) "Just fill in this form – A review for designers." *Applied Ergonomics*, 6:213-220.

Wright, P. and P. Barnard. (1978) "Asking Multiple Questions about Several Items: The Design of Matrix Structures on Application Forms." *Applied Ergonomics*, 9:7-14.

**Figure 1: Question formats used to test hypotheses of symbolic and graphical changes in questionnaire formats.**

## Graphical Language Changes

*Location of special instruction before vs. after answer categories.*

A. Have one-on-one meetings with professors contributed significantly to your WSU education?

If you haven't had many one-on-one meetings, just skip to Question 9.

☐ Yes
☐ No

B. Have one-on-one meetings with professors contributed significantly to your WSU education?

☐ Yes
☐ No

If you haven't had many one-on-one meetings, just skip to Question 9.

---

*Linear vs. nonlinear scale layout. First of 2 experimental items.* [1]

A. Overall, how would you rate the quality of education that you are getting at WSU?

☐ Excellent ☐ Good ☐ Poor
☐ Very good ☐ Fair

B. Overall, how would you rate the quality of education that you are getting at WSU?

☐ Excellent
☐ Very good
☐ Good
☐ Fair
☐ Poor

---

*Equal vs. unequal spacing between response boxes. First of 2 experimental items.* [1]

A. What percentage of WSU students do you think use the Student Recreation Center? [2]

☐ Less than one-fourth of WSU students
☐ About half of WSU students
☐ About three-fourths of WSU students
☐ More than three-fourths of WSU students
☐ No opinion

B. What percentage of WSU students do you think use the Student Recreation Center? [2]

☐ Less than one-fourth of WSU students
☐ About half of WSU students
☐ About three-fourths of WSU students
☐ More than three-fourths of WSU students
☐ No opinion

---

**Figure 1: Question formats used to test hypotheses of symbolic and graphical changes in questionnaire formats. (Continued)**

**Graphical language changes**

*Use of larger answer spaces on open-ended answers. First of 3 experimental items.*[1]

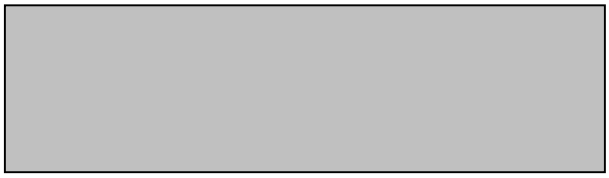
A. Why did you choose to attend Washington State University?

B. Why did you choose to attend Washington State University?

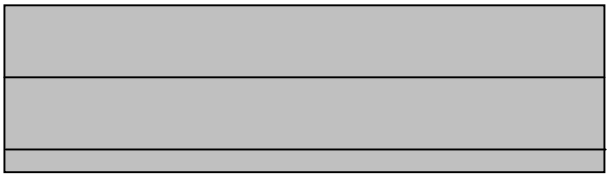
**Symbolic Language Changes**

*Addition of lines in open-ended questions. First of 2 experimental items.*[1]

A. What store(s) or business(es) would you most like to see in the Pullman and Moscow area that are currently not available?



B. What store(s) or business(es) would you most like to see in the Pullman and Moscow area that are currently not available?



*Addition of arrow to identify subordinate question*

A. After finishing school at Washington State University, where do you hope to live?

☐ Eastern Washington
☐ Somewhere else ➔ Where?_____

B. After finishing school at Washington State University, where do you hope to live?

☐ Eastern Washington
☐ Somewhere else Where?_____

**Figure 1: Question formats used to test hypotheses of symbolic and graphical changes in questionnaire formats. (Continued)**

---

**Graphical and Symbolic language change**

*Polar point scale vs. answer box. First of 3 experimental items.*[1]

A. On a scale of 1 to 5, with one being very satisfied and 5 being very dissatisfied, how satisfied are you with the classes you are taking this semester?

- ☐ 1 Very Satisfied
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 Very Dissatisfied

B. On a scale of 1 to 5, with one being very satisfied and 5 being very dissatisfied, how satisfied are you with the classes you are taking this semester?

_____ Number of your rating

---

[1]Wording and visual layout of other items is available from the authors.
[2]The query was displayed on one continuous line.

**TABLE 1: INDEPENDENT GRAPHICAL LANGUAGE MANIPULATIONS**

**Graphical location**

Percentage of respondents choosing each response when instruction to skip located before vs. after the response options

| Responses | Instructions before | Instructions after |
|---|---|---|
| **(n)** | **519** | **523** |
| Yes | 54.7 | 54.9 |
| No | 19.1 | 40.3 |
| Missing | 26.2 | 4.8 |
| **Total** | 100% | 100% |

Answered vs. Missing - $\chi^2 = 106.67$ p=.000
Yes vs. No vs. Missing - $\chi^2 = 116.99$ p=.000

**Linear vs. Nonlinear scale layout**

Percentage of respondents choosing each response for linear vs. nonlinear graphical layout

| | Question 1 | | | Question 2 | |
|---|---|---|---|---|---|
| **Responses** | **Linear** | **Nonlinear** | | **Linear** | **Nonlinear** |
| **(n)** | **518** | **517** | | **519** | **518** |
| (1) Excellent | 11.4 | 11.0 | (1) Very Satisfied | 42.0 | 38.8 |
| (2) Very Good | 48.8 | 42.4 | (2) Somewhat Satisfied | 49.5 | 51.2 |
| (3) Good | 31.3 | 40.4 | (3) Somewhat Dissatisfied | 6.9 | 8.9 |
| (4) Fair | 7.0 | 5.4 | (4) Very Dissatisfied | 1.5 | 1.2 |
| (5) Poor | 1.5 | .8 | | | |
| **Total** | 100 | 100 | | 100 | 100 |
| **Mean** | 2.4 | 2.4 | | 1.7 | 1.7 |
| Difference of Means - t=.8 p=.206 | | | Difference of Means – t=1.1 p=.146 | | |
| Overall chi-square - $\chi^2 = 10.8$ p=.029 | | | Overall chi-square – $\chi^2 = 2.3$ p=.509 | | |
| Good vs. Very Good - $\chi^2 = 8.2$ p=.004 | | | Top line (1,3) vs. Bottom (2, 4) - $\chi^2 = .16$ p=.685 | | |
| Top line (1,3, 5) vs. Bottom (2, 4) - $\chi^2 = 6.7$ p=.010 | | | No significant differences for individual responses | | |

**TABLE 1: INDEPENDENT GRAPHICAL LANGUAGE MANIPULATIONS (CONT.)**

**Equal vs. unequal spacing between response categories**

Percentage of respondents choosing each response with equal vs. unequal spacing between categories

| Responses | Nominal scale Equal | Nominal scale Unequal | | Ordinal scale Equal | Ordinal scale Unequal |
|---|---|---|---|---|---|
| **(n)** | **474** | **471** | | **510** | **514** |
| To have a life partner with whom you have a satisfying relationship[1] | 30.8 | 37.6 | (1) Less than one-fourth of WSU students[1] | 17.3 | 19.1 |
| Enjoy your work | 43.9 | 42.9 | (2) About half of WSU students | 47.5 | 46.5 |
| Earn a high income | 10.6 | 7.9 | (3) About three-fourths of WSU students | 24.7 | 22.6 |
| Raise a family | 14.8 | 11.7 | (4) More than three-fourths of WSU students | 4.3 | 4.7 |
| | | | No opinion[1] | 6.3 | 7.2 |
| **Total** | 100% | 100% | | 100% | 100% |
| | | | **Mean** | 2.3 | 2.3 |
| Overall chi-square - $\chi^2 = 6.8$ p=.079 | | | Difference of Means – t=.07 p=.472 | | |
| To have a … vs. all others - $\chi^2 = 4.8$ p=.028 | | | Overall chi-square - $\chi^2 = 1.4$ p=.844 | | |
| | | | No significant differences for response option 1 or No opinion when run separately or together | | |

[1] These answer boxes were separated by one or two lines from the other responses

**Use of larger spaces on open-ended questions**

Mean number of words and themes reported by respondents in larger vs. smaller answer spaces

| Question | Mean # of | Larger space | Smaller Space | Difference of Means | |
|---|---|---|---|---|---|
| Why attend WSU | **(n)** | **500** | **509** | | |
| | Words | 13.3 | 9.7 | t=6.5 | p=.000 |
| | Themes | 2.0 | 1.8 | t=2.7 | p=.003 |
| Description of advisor | **(n)** | **484** | **494** | | |
| | Words | 12.9 | 6.6 | t=12.9 | p=.000 |
| | Themes | 2.1 | 1.7 | t=8.0 | p=.000 |
| Add'l recreational activities | **(n)** | **284** | **280** | | |
| | Words | 12.0 | 10.2 | t=1.8 | p=.039 |
| | Themes | 1.5 | 1.4 | t=0.7 | p=.195 |

**TABLE 2: INDEPENDENT SYMBOLIC LANGUAGE MANIPULATIONS**

**Addition of lines on open-ended answer spaces**

Mean number of words and themes reported by respondents with lines vs. no lines

| **Question** | **Mean # of** | **Lines** | **No Lines** | **Difference of Means** | |
|---|---|---|---|---|---|
| Add'l stores in Pullman | **(n)** | **401** | **407** | | |
| *(Smaller answer space)* | Words | 5.3 | 6.7 | t=-4.9 | p=.000 |
| | Themes | 2.3 | 2.4 | t=-.6 | p=.269 |
| Changes for WSU | **(n)** | **371** | **394** | | |
| *(Larger answer space)* | Words | 27.9 | 28.8 | t=-.5 | p=.296 |
| | Themes | 1.7 | 1.7 | t=-.3 | p=.391 |

Only 5 respondents (1.2%) on the stores question and 3 respondents (.8%) on the changes question did not use the lines when provided

Respondents who left the changes for WSU question blank was significantly larger on the lines version (26.6% vs. 19.4%, $\chi^2 = 7.75$ p=.005).

**Addition of arrow to identify subordinate question**

Percentage of respondents answering subordinate question

| | **Arrow** | **No Arrow** |
|---|---|---|
| **Total mentions (n)** | **429** | **434** |
| Mentions from Eligible | 93.9 | 90.8 |
| Mentions from Noneligible | 2.1 | .5 |
| Missing (No response) | 4.2 | 8.8 |
| **Total** | 100% | 100% |

Total mentions vs. Missing - $\chi^2 = 7.34$ p=.007

Eligible vs. Noneligible mentions vs. Missing - $\chi^2 = 11.65$ p=.003

Eligible mentions vs. Noneligible mentions - $\chi^2 = 4.26$ p=.039

**TABLE 3: COMBINED GRAPHICAL AND SYMBOLIC LANGUAGE MANIPULATION**

**Polar point scale vs. answer box**

Percentage of respondents choosing each response on polar point vs. answer box

| | Question 1 | | Question 2 | | Question 3[1] | |
|---|---|---|---|---|---|---|
| **Responses** | **Polar Point** | **Answer box** | **Polar Point** | **Answer box** | **Polar Point** | **Answer Box** |
| **(n)** | **517** | **513** | **517** | **512** | **506** | **466** |
| (1) Very Satisfied | 15.9 | 9.8 | 10.8 | 6.3 | 15.2 | 14.2 |
| (2) | 43.5 | 29.2 | 47.6 | 35.4 | 31.8 | 23.8 |
| (3) | 31.0 | 34.5 | 31.3 | 33.8 | 31.6 | 29.6 |
| (4) | 7.7 | 21.3 | 9.5 | 20.9 | 15.4 | 21.7 |
| (5) Very Dissatisfied | 1.9 | 5.3 | 0.8 | 3.7 | 5.9 | 10.7 |
| **Total** | 100% | 100% | 100% | 100% | 100% | 100% |
| **Mean** | 2.4 | 2.8 | 2.4 | 2.8 | 2.7 | 2.9 |
| **Difference of Means** | t=-7.7 p=.000 | | t=-6.9 p=.000 | | t=-3.5 p=.000 | |
| **Chi-square** | $\chi^2 = 63.4$ p=.000 | | $\chi^2 = 48.1$ p=.000 | | $\chi^2 = 18.0$ p=.001 | |

[1] Response categories to this item were (1) Outstanding to (5) Terrible